A Hybrid Certificate Management for Mobile ad-hoc Networks

Mrs. Hemalatha Jai Kumari1 , Dr. A. Kannammal2

Bharathiar University, India

Abstract:
Mobile adhoc network (MANET) applications are gaining importance due to increased number of personal devices and ubiquitous computing. Authenticity is the most fundamental issue in these applications, since a breach of authenticity leads to a system wide compromise. The existing public key infrastructure (PKI) handles the applications in a wired network using a centralized certificate server. This server handles the creation, renewal and revocation of certificates. This procedure is however impossible in mobile adhoc networks due to the absence of fixed infrastructure and centralized management. Apart from these, dynamic topology and link failure may also result in re-authentication and may warrant for timely communication. Most of the existing protocols are based on pre-shared secret or pre-obtained public key certificates. This assumption has some practical weaknesses for some emergency applications using MANET. To overcome the shortcomings, a wide covered network is used to design a secure certificate distribution scheme. The protocol steps are also discussed for this scheme and it is found to be efficient against security attacks and suitable for MANET. The computational cost is given and it is found to be reasonable.

Keywords:
MANET, PKI, authentication, threshold cryptography, security.

1. INTRODUCTION

A Mobile Adhoc Network (MANET) is a collection of wireless nodes that can dynamically form a network to exchange information without using any pre-existing fixed network infrastructure. There should be a field of pervasive environment, which facilitates the communication between the mobile devices. Thus, a new model of pervasive computing including new architectures, standards, devices, services, tools and protocols must be developed for MANET. Attractive applications of MANET include Military battlefield, commercial sector, personal area network (PAN) that is forming a temporary network with devices with mobility and local level, that is providing the link at an instant using notebook computer and palmtop computers and other civilian environments.

Regardless of the attractive applications, the features of MANET include several challenges. Security is one of the vital issues in MANET. Since adhoc networks rely on wireless communication medium, it is important to deploy a security protocol to protect the privacy of transmissions. Due to open network architecture, shared wireless medium and dynamic network topology, providing security services among the nodes in MANET is an important issue and a non-trivial challenge [9]. However, the implementation schemes of key management, authentication and authorization are different because there is no aid of a trusted third-party certification authority to create trusted relationships by exchanging private/public key pairs [13].

Several research studies [2-8] have focused mainly on a secure routing protocol or secure group communication for MANET. They assumed that all nodes have pre-shared common secret or public-key certificates obtained before joining the network. This assumption has a practical weakness for some emergency applications because some nodes without pre-obtained certificates will be unable to join the network.

Mobile ad-hoc network applications are increasing due to the tremendous growth of personal devices and ubiquitous techniques. Hence, the number of users of a network may increase and consequently the number of users without certificates or expired certificates may increase. For remedying this practical weakness, these nodes may join the network and they can use the proposed scheme to obtain valid certificates.

The paper is organized as follows: Section 2 discusses the reviews and related work. Section 3 discusses the proposed system design and the notations used along with the protocol steps. Section 4 gives the conclusion.

2. REVIEW AND DISCUSSIONS ON RELATED WORK

Several famous solutions [10-13] for public-key management for mobile adhoc networks are available. Following is a brief review and discussion about the pros and cons of some of the significant contributions.

2.1 Threshold Public-Key Management with Partially Distributed Authority

Zhou and Haas [13] proposed a threshold public-key management scheme. This scheme used a pair of system public/secret keys. They assumed that the public key is

Manuscript received June 5, 2012
Manuscript revised June 20, 2012

known for all nodes and the secret key is divided into ‘n’ shares (where ‘n’ is the number of nodes), using a threshold cryptography scheme. If a node can obtain some minimum (threshold) number ‘t’ of partial certificates, the requesting node can combine them to produce one complete certificate. However, the approach raises three non-trivial issues:

Issue #1: The policy of how to choose ‘n’ specific nodes?

Issue #2: When a node’s public key needs to be signed, whether there is a threshold number for the ‘n’ chosen specific nodes available to the requesting node?

Issue #3: The third issue is the decision about the threshold value, which is a trade-off between availability and robustness. If the threshold value is large, the availability will be decreased but, this increases the robustness.

2.2 Threshold Public-Key Management with Fully Distributed Authority

To improve the fairness and availability of the Zhou-Hass scheme, Kong et al. [12] proposed a fair scheme. It is based on threshold cryptography and shared secrets. The basic goal of a threshold secret sharing method is to share a secret key ‘k’ among an arbitrarily large community using a secret polynomial f(x). If the degree of f(x) is (k-1), any ‘k’ members of the community can recover the secret key, while any number of members less than ‘k’ reveals no information of the secret [13]. Based on this, a node receives its public key from its ‘k’ neighboring nodes. Here, ‘k’ is a parameter, which needs to be carefully tuned so that the method is effective.

The certificate creation process is as follows: Initially all the nodes in the network used to be bootstrapped with their certificates from a trusted central management. When a new node wants to obtain its certificate, it sends a request to its ‘k’ neighboring nodes requesting for partial certificates, which are then combined together by the target node to issue the new certificate using an interpolation function. The drawback is an attacker could take as many identities as necessary to collect enough shares, and thereby construct the secret key.

2.3 Threshold ID-Based Public-key Management

In 2004, Deng and Agrawal [11] proposed a threshold ID-based public key management scheme. In their system, all initial nodes compute collaboratively a pair of public/private keys without constructing the system private key at any single node. They assume that each node has a unique IP address or identity when it joins the network. Each node within the network obtains its partial system private key corresponding to its identity. Since each node in the network shares a part of the system private key, each node can be held responsible for generating other nodes’ personal private keys based on a threshold. In the initialization phase, each node must contact at least ‘t’ neighbor nodes and present its identity and requests for a personal private key. The ‘t’ nodes work together to generate the personal private key corresponding to the presented identity. In order to ensure that the generated shares are securely transmitted, the requested node must present its self-generated public key when sending the request message. The ‘t’ issuing nodes send encrypted shares to the requesting node using its temporary public key which is considered as a drawback.

2.4 Self-Organized Public-Key Management

In 2003, Capkun [10] proposed a fully self-organized scheme that required no trusted authority. One of the certificate-based authentication methods proposed is based on the formation of certificate graphs. The suggested method is an extension of PGP (Pretty Good Privacy) certificates. Each node maintains an updated and non-updated local certificate repository, which consists of subset of updated expired certificates. The use of two repositories is required in providing a good estimate of the certificate graph and for node authentication.

The drawbacks of this mechanism are: i. expensive table that has to be maintained for the certificate repositories and ii. Each time a node moves from one locality to another, it must regenerate table with other nodes and update the tables again. Obviously, this approach is computationally expensive.

In 2004, Varadharajan [7] discussed the security of a mobile adhoc network. They suggest that each node must obtain a valid certificate before joining the mobile adhoc network. However, this assumption has a practical weakness for some emergency applications, because some nodes do not pre-register to the system authority before joining the network. Therefore, they cannot obtain valid certificates.

Security for a mobile adhoc networks is a recent problem and some of the existing solutions are discussed. It can be observed that no single solution is the best solution and they have their own limitations. Further, the existing schemes cannot accommodate the frequent membership change and the dynamism. Dynamic membership behavior is a common phenomenon in MANET and a solution has been provided for this problem for authentication using a wide covered network. The proposed scheme is elaborately discussed in the next section.

3. PROPOSED SYSTEM DESIGN

In a mobile adhoc network, there are many nodes represented as in Fig 1. Each dot denotes mobile nodes (mobile devices). The network route from sender node to the receiver node could require a number of intermediate nodes to forward the packets. Each node forwards the packets to its neighboring nodes. Assume that all wireless transmission links in this network are bidirectional. There are two kinds of nodes. First node is a general node (denoted by black dots) with an adhoc network card. A general node is a legal user of some server in the Internet. It owns an account/password pair for accessing the corresponding server. General node may send messages to its corresponding server in the Internet through the communication channel constructed by an agent node. Second node is an agent node (denoted by double circle) that possesses both adhoc card and a network card. The agent node can communicate with the neighborhood nodes and other wide covered networks (WCN) (such as satellite or cellular networks). The wide covered network can then connect to the Internet and its service area is fully covered by a large place. There exists a certificate authority (CA) for issuing certificates. The authentication server of the wide-covered network has a certificate issued by the CA.

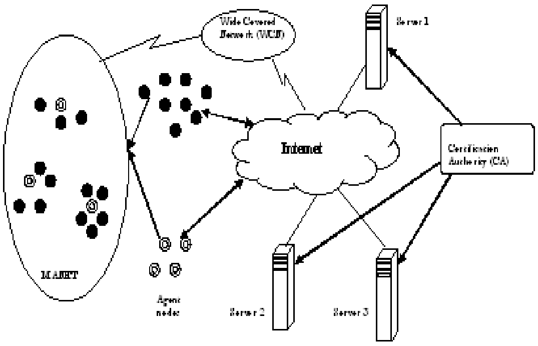


Fig. 1 Proposed System Design

3.1 Notations used

Table 1 gives the details of the system parameters and the notations used in the protocol.

3.2 Procedure Description

The protocol is primarily designed for emergency and rescue operations. In such scenarios, very dynamic and unpredictable situations are expected. New nodes frequently appear and others disappear causing frequent network partitioning and merging. An important characteristic of an emergency and rescue scenario is that the organization involved (police, fire department, paramedics, etc) are often well-structured, public entities. Some of them might have sensitive data on the scene, like medical or police records that are highly confidential and should remain such. Before the rescue personnel of the different organizations comes to the rescue scene, all devices are prepared for their tasks. One task in the preparation phase, which is called a *priori* phase, is the installation of valid certificates. The certificates are signed by a commonly trusted authority, such as the ministry of internal affairs, ministry of defense etc., which is on the top of trust chain. Nodes can possibly authenticate each other without need for contacting a third party. Therefore, there is no need for a fully self-organized public key management system that does not rely on trusted authorities. The nodes that have the possibility to authenticate each other without need for contacting a third party are termed as **Agent nodes**. They are the leaders of the group. The Agent nodes are given two cards: **adhoc card and network card**. The other nodes are their subordinates. They are called **general nodes.** They posses only the network card.

a. ***SK_A, PK_A – Secret Key of Entity A and its corresponding public key.***

b. ***R_A – A large random value generated by entity A.***

c. ***h ( ) – a one-way hash function.***

d. ***E_{PK_A} ( ), D_{SK_A} ( ) – data encryption with public key PK_A, and the corresponding data decryption with secret key SK_A.***

e. ***$Cert_A$ – the certificate of entity A.***

f. ***Sig $_{SK_A}$ ( ) – a digital signature generated by entity A using secret key SKA.***

g. ***PW_{ID} – a password or symmetric secret shared by a user with identity ID and the corresponding server.***

h. ***$Chain_{A\text{-}B}$ – a certificate chain from entity A to the highest entity B.***

Table:1 System parameters and notations used

As the agent nodes possess adhoc network card, it establishes the communication channel with the wide-covered network. General nodes can communicate with the server only through the communication channel established by the agent node.

As the communication takes place only through the agent nodes, if the agent node leaves the network, the general nodes cannot communicate with the wide covered network or Internet. In order to avoid this, whenever an agent node leaves the network, it uses **delegation or authentication forwarding** mechanism to authenticate a general node to become the new agent for the group. This formulates a cluster. The new agent node is selected according to the time of availability in the cluster. More the general node stays; it becomes the next eligible node to become as an agent node.

The entire mobile adhoc network is divided into a number of clusters. Each cluster has a cluster head CH (agent node). Nodes with dual interface network cards are identified as cluster head. All the nodes belong to exactly one cluster. Cluster construction is done in two phases: Initial Phase and Certificate Distribution Phase.

Each agent node i.e., cluster head in the mobile adhoc network can construct a communication channel with the wide-covered network (WCN). Other nodes in the same cluster may use the communication channel constructed by the cluster head for accessing the Internet. The nodes without certificates in the mobile adhoc network must be either agent nodes or general nodes.

If a node without a certificate in the mobile adhoc network is an **agent node**, it can directly access the wide covered network and use the proposed scheme to obtain a certificate issued by the authentication server of the wide covered network.

If a node without a certificate in the mobile adhoc network is a **general node**, then it can access the corresponding server on the Internet through the communication channel created by one agent node. It also uses the proposed scheme to obtain the certificate issued by the corresponding server. Hence the important part of the discussion is the cluster creation.

3.3 Cluster Creation

Each node with dual-interface network cards is a cluster head (CH). Without loss of generality, it is assumed that the mobile adhoc network includes ‘n’ nodes and there are ‘m’ agent nodes; thus the mobile adhoc network will be divided into ‘m’ disjoining clusters. According to the value of ‘m’, there are three possible cases:

Case 1: If m = n, each node in the mobile ad hoc network is one agent node.

Case 2: If m = 1, only one node is an agent node and all other nodes must use this node to get certificates.

Case 3: If 1 < m < n, there are many clusters. For example, the number of clusters in Figure 1 is four.

The design concept of the cluster creation is:

First, the cluster heads (CHs) must broadcast their connecting information (CI) to their neighborhoods, where CI includes the identity of CH and the hop distance from the CH. If a node receives only one CI, then the node sends a join message to the corresponding CH. If a node receives many CI’s, the node then may choose one CH and sends a join message to the corresponding CH.

Second, each node that has joined some CH’s also broadcasts the connecting information (CI) to its neighborhoods. Other non-joined nodes may choose one CI with the smallest hop distance or according to other weight information to join the corresponding CH. The weight information may include computation ability and the hop count scope of each node.

3.4 Protocol steps

In this protocol, there are two components: a cluster head (CH) and the authentication server of the wide covered network (WCN-AS). Assume that the WCN-AS with identity C has a secret key SK_C and the corresponding public key PK_C. The protocol allows an agent node i.e., Cluster Head (CH) to use the wide covered network to obtain the certificate. Table 2 gives the steps involved in the algorithm.

1. ***If the cluster head is already attached to the wide-covered network, go to step 2. Else, the cluster head (CH) accesses the wide-covered network and gets the authentication.***
2. ***If the cluster head leaves the network for some reasons, it selects a general node that stays for a longer period in the network and delegates the general node as a cluster head.***
3. ***Let HID denote the identity of the cluster head.***
4. ***The Cluster head (CH) with identity HID randomly selects a secret key SK_{HID} and computes the corresponding public key PK_{HID}.***
5. ***The CH also chooses a random number R_{HID} and uses the PK_C of the WCN-AS to encrypt R_{HID} and HID.***
6. ***The CH then computes***
   $hv = h\ (PW_{HID}, R_{HID}, HID, PK_{HID})$.
7. ***The CH sends the requesting message (HID, PK_{HID}, E_{PK_C} (HID || R_{HID}), hv) to the WCN-AS.***
8. ***Upon receiving the request message, the WCN-AS first performs***
   D_{SK_C} (E_{PK_C}, (HID || R_{HID})) to recover R_{HID} and then validates whether
   $hv= h\ (PW_{HID}, R_{HID}, HID, PK_{HID})$ holds or not.
9. ***If the check is correct, the WCN-AS generates the certificates***
   $Cert_{HID} = (PK_C, HID, PK_{HID}, T, Sig_{SK_C}\ (HID, PK_{HID}, T))$,where T is the valid time period.
10. ***The WCN-AS then sends (HID, $Cert_{HID}$, PK_C, $Cert_C$, $Chain_{C\text{-}CA}$) to the CH.***

Table: 2 Authentication Algorithm

Finally, CH obtains the certificate $Cert_{HID}$ and then validates it. Since the certificate $Cert_{HID}$ is issued by the WCN-AS, and other nodes cannot attempt to validate $Cert_{HID}$ as they do not know PK_C. Therefore, CH may validate $Cert_C$ and $Cert_{C\text{-}CA}$ before validating $Cert_{HID}$. To protect the password, both CH and WCN-AS may update PW_{HID} with R_{HID}. This is called as one time password. When the CH accesses the WCN-AS the next-time the CH has to use the new password to computer the hash value of ‘h’.

3.5 Validation against Security Attacks

The proposed design is validated for some of the standard security attacks.

Passive attack: If a malicious attacker intercepts a valid request message (HID, PK_{HID}, N, E_{PK_C} (HID || R_{HID}), hv) where h = H (PW_{HID}, HID, PK_{HID}), he cannot compute the CH's secret key SK_{HID} from PK_{HID} because this encryption is based on the public-key based encryption system E_{PK_A} ( ) / D_{SK_A} ( ). Since PW_{ID} is protected under the one-way hash function hv = H (PW_{HID}, R_{HID}, HID, PK_{HID}), the attacker cannot obtain PW_{ID} from it. Therefore, proposed protocol can withstand passive attacks.

Active attack: Suppose that an attacker tries to guess the password PW_{HID} of the CH with identity HID. The attacker must use all possible candidate passwords to compute all possible hash values and compare them with hv = H (PW_{HID}, R_{HID}, HID, PK_{HID}). However, the large random number R_{HID} is encrypted using public key PK_C of the WCN-AS. Thus, only the WCN-AS with secret key SK_C can decrypt it. Obviously, the guessing attack cannot work. Therefore, the proposed protocol can withstand guessing attacks.

4. PERFORMANCE METRICS

To evaluate the effectiveness of the proposed protocol the following metrics are used:

★ Network Size

It is the number of nodes in a cluster. The performance of the proposed protocol is greatly affected by the number of nodes present in the network and therefore it is taken as a test parameter for the evaluation of the software.

★ Cluster size

The performance of any algorithm in mobile ad hoc network is greatly affected by the cluster size formed. It is found that, many algorithms fail when the number of nodes in a cluster increases or decreases. So it is decided to test the algorithm with respect to the number of nodes in a cluster.

★ Number of clusters

The key generation and reply time changes with the number of clusters in the ad hoc network. As this parameter influences the functioning of the algorithm, it is decided to use number of cluster as a performance metric.

★ Clustering Time

Since the topology in ad hoc network changes dynamically and frequently, the protocol execution performance is closely related to the clustering time and hence is used for analyzing the performance.

★ Coverage Range

The formation of cluster heads, clusters and generation of keys are greatly influenced by the coverage or transmission range and therefore coverage range is also taken as a performance evaluation parameter.

★ Key length

Key length is a term that indicates the length of the secret key generated. The security of the proposed system is dependent upon the key lengths being used and is, for that reason, used for performance evaluation.

★ Reply Time

When a node joins the network for communication, it sends a request for communication with its certificate to the server. The result can be either a successful join operation (valid certificate) or reject operation (invalid certificate). Reply time is the time taken for the joining node to get a successful reply from the server. Since a node is permitted to communicate only when it has a valid certificate, the reply time is considered as an important parameter for performance evaluation.

★ Secret Key Generation Time

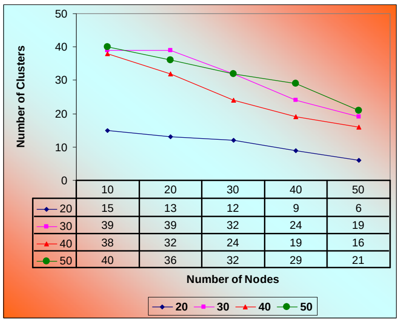
Key generation is the process of generating Secret keys for mobile nodes to use for certification. The time required to calculate the secret key on each visited host greatly affects the working of the system and hence it is considered.

★ Number of Hops

When each new node joins a cluster head, it broadcasts connecting information to its neighboring nodes. The connecting information has details like weight information, computability ability, hop count, etc. Other non-joined nodes may choose one CI with the smallest hop distance. As finding an optimal node is important parameter, it is considered as a performance metric in the evaluation.

5. PERFORMANCE ANALYSIS

Performance of the protocol with respect to network size and number of clusters formed varying the coverage range is analyzed in this section. The simulated environment had 50 nodes, that is, the maximum number of nodes in the network is in the range 20-50 and the coverage area varied from 100 to 500 km.



	10	20	30	40	50
◆ 20	15	13	12	9	6
■ 30	39	39	32	24	19
▲ 40	38	32	24	19	16
● 50	40	36	32	29	21

Fig 5.1 Network size and number of clusters over varying coverage area.

It is evident from Fig. 4.1, that the number of cluster is inversely proportional to the number of nodes in the network. While varying the coverage area of the ad hoc network, this trend is noticed to be more predominant when the network size becomes larger.

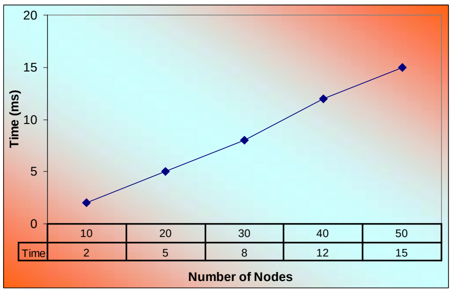


Fig 5.2 Network size and clustering time

Clustering time is affected to a great extent by the network size. It's been observed that when more and more mobile nodes join the network, clustering time also increases irrespective of the coverage area.

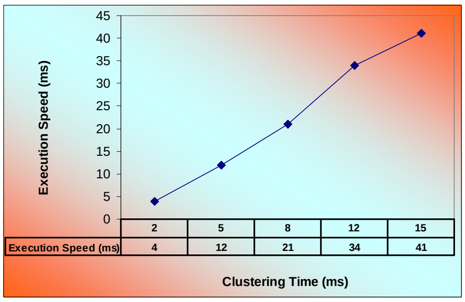


Fig 5.3 Clustering time Vs Execution Speed

The increase in clustering time directly affects the execution speed of the protocol. This is because, when the number of nodes is increased, the size of the repository overcomes the storage capacity. If only a few certificates are stored, then there might exist a performance problem which may affect the network

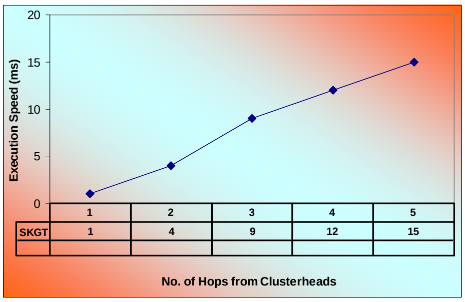


Fig 5.4 Hop length and secret key generation time

The result reflects that keeping the hop length to a minimum also reduces the communication overhead but is challenging because of the changing topology of the mobile architecture.

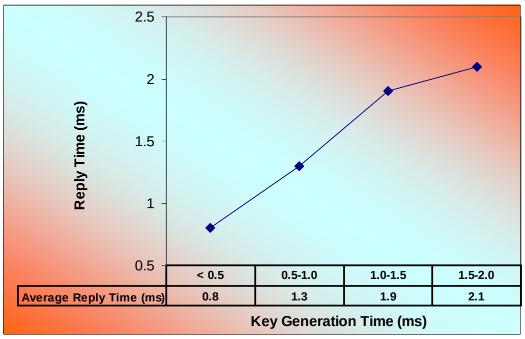


Fig. 5.5: Key Generation Time and reply time

It's been observed that the key generation time increases because of various reasons like transmission range, network size, clustering time, key size, etc.

6. CONCLUSION

In this paper, an efficient authentication scheme using public key infrastructure for highly dynamic adhoc networks is presented. The protocol provides concrete identity information on nodes. The nodes without pre-obtained certificates in the mobile adhoc networks must be either agent nodes or general nodes. If the node is an agent node, it uses the given Protocol to obtain a certificate issued by the authentication server of the wide covered network. The general node has to be authenticated by the cluster head and it can use the above protocol to obtain a valid certificate, as they are the legal subscribers of the wide covered network or legal users of some server on the Internet. The computational cost is also very reasonable in this case. Therefore, the proposed scheme provides concrete node identity confirmation and found to be efficient in handling security attacks.

REFERENCES

[1] Charlie Kaufman, Radia Perlman, Mike Speciner, Network Security: PRIVATE Communication in a PUBLIC World, II Edition, Prentice Hall 2005.

[2] Challal Y, Bettahar H, Bouabdallah A. SAKM: a scalable and adaptive key management approach for multicast communication. ACM SIGMOBILE Computer Communications Review 2004; 34(2): 55–70.

[3] Gupte S, Singhal M. Secure routing in mobile wireless ad hoc networks. Ad Hoc Networks 2003; 1(1): 151–174.

[4] Hu YC, Johnson DB, Perrig A. SEAD: secure efficient distance vector routing for mobile wireless ad hoc networks. Ad Hoc Networks 2003; 1(1): 175–192.

[5] Li XY, Wang Y, Frieder O. Efficient hybrid key agreement protocol for wireless ad hoc networks. In Proceedings of the Eleventh International Conference on Computer Communications and Networks, 2002; 404–409.

[6] Tseng YM, Yang CC, Liao DR. A secure group communication protocol for ad hoc wireless networks. In 10th Mobile Computing Workshop (MCW 2004), Taiwan, 2004; 267–276.

[7] Varadharajan V, Shankaran R, Hitchens M. Security for cluster based ad hoc networks. Computer Communications 2004; 27(5): 488–501.

[8] Zapata MG. Secure ad hoc on-demand distance vector (SAODV) routing. ACM SIGMOBILE Mobile Computing and Communications Review 2002; 6(3): 106–107.

[9] Yang H, Luo H. Y, Ye. F., Lu. S. W., and Zhang. I., “Security in Mobile Ad hoc Networks: Challenges and Solutions,” IEEE wireless communications 2004, 11 (1), pp 38-47..

[10] P. Caballero-Gil and C. Hern´andez-Goya, “Self-organized authentication in mobile ad-hoc networks,” Journal of Communications *and Networks*, vol. 11, no. 5, pp. 509–517, 2009.

[11] Deng H, Agrawal DP. TIDS: threshold and identity-based security scheme for wireless ad hoc networks. Ad Hoc Networks 2004; 2(3): 291–307.

[12] Kong J, Zerfos P, Luo H, Lu S, Zhang L. Providing robust and ubiquitous security support for mobile ad hoc networks. In IEEE Ninth International Conference on Network Protocols (ICNP’01), November 2001; 251–260.

[13] Zhou. L. and Haas. Z. J., “Securing ad hoc networks,” IEEE Network Journal, 1999, 13 (6), pp 24-30..

ABOUT AUTHORS

1. Mrs. Hemalatha Jai Kumari , is a part-time research scholar at Bharathiar University. Her area of interest is security and privacy in ad hoc networks.
2. Dr. A. Kannammal, is an associate professor, department of computer applications, Coimbatore Institute of Technology, Coimbatore. Her area of interest is Web security and e-commerce applications.